数理解析研究所講究録
第 1584 巻 2008 年 1-7

対称錐上の主双対内点法と双対幾何構造
Primal-Dual rierior-Pori Methods and Dualistic Structure on Symmetric Cones

東北学院大学工学部機械知能工学科　魚橋 慶子（Keiko UOHASHI）
uohashi@tjcc.tohoku-gakuin.ac.jp;　Department of Mechanical Engineering
& Intelligent Systems, Faculty of Engineering, Tohoku Gakuin University

1．序

対称錐上の線形計画法は，線形計画法（ＬＰ）・半正定値計画法（ＳＤＰ）を含むクラスであるという理由から，または対称錐計画法自体への興味から，近年研究が進められている．線形計画法・半正定値計画法の解法として発展した主双対内点法は，対称錐上へも拡張されている．しかし線形計画法・半正定値計画法における場合と異なり，対称錐上における主双対内点法と双対幾何（情報幾何）との関係は明らかでない．

そこで本研究では，対称錐線形計画法の主双対内点法に現れる概念を双対幾何の概念へ対応させることを試みる．まずユークリッド的ジョルダン代数と対称錐について概説する．次に対称錐上の最適化問題について，線形制約をもつ場合に焦点を当て説明する．そして双対幾何と主双対内点法との関連を述べる．最後に最適化におけるスケーリングと作用素平均について述べる．

本稿作成にあたり，微分幾何を専門としない最適化研究者と，最適化を専門としない微分幾何研究者の双方に理解し易いよう心がけた．そのため詳しい説明に欠ける箇所や，逆に説明過多の箇所が見られる．説明不足を補うためには参考文献をご覧願う．

2．ユークリッド的ジョルダン代数と対称錐

ジョルダン代数と対称錐に関しては文献 [7] が詳しい．本章では我々の研究に必要な事項をまとめる．

ベクトル空間 V は，次をみたす積 $*$ が定義されているとき，ジョルダン代数（Jordan algebra）とよばれる．

$$
\begin{aligned}
x*y &= y*x, \\
x*(x^2*y) &= x^2*(x*y)
\end{aligned}
$$

（$x^2=x*x$）for $x,y\in V$．以下 V を，単位元を持つ $\mathbf{R}$ 上の n 次元ジョルダン代数とする．ここで $e\in V$ が単位元（identity element）であるとは，$x*e=e*x=x$，$x\in V$ をみたすことをいう．また $x\in V$ に対し $x*y=e$ をみたす $y\in\mathbf{R}[x]$ が存在するとき，x は可逆である（invertible）という（$\mathbf{R}[x]$ は $\mathbf{R}$ 上における x の多項式環）．$\mathbf{R}[x]$ は結合的（associative）であるから，$y\in\mathbf{R}[x]$ は一意に定まる．ゆえに $x*y=e$ をみたす y は x の逆元（inverse）とよばれ $x^{-1}=y$ と記される．

ジョルダン代数 V が単純である（simple）とは，V が非自明なイデアルを持たないこ

とをいう．またジョルダン代数 V がユークリッド的である（Euclidean）とは，V が結合的な正定値対称双線形 2 次形式を持つことをいう．ここで 2 次形式 $\langle\, ,\rangle$ が結合的であるとは，任意の $x, y, z \in V$ に対し $\langle L(x)y, z\rangle = \langle y, L(x)z\rangle$ を満たすことをいう．実際には，同値な性質である形式的に実（formally real）（$x^2 + y^2 = 0$ ならば $x = 0, y = 0$）という性質がしばしば利用される．元 $x \in V$ に対し，$L(x)$ と $P(x)$ をそれぞれ次で定義される V の自己準同形（endomorphism）とする．$P(x)$ は V の 2 次表現（quadratic representation）とよばれる．

$$
\begin{aligned}
L(x)y &= x * y,\ y \in V \\
P(x) &= 2L(x)^2 - L(x^2).
\end{aligned}
$$

次に対称錐について説明する．集合 Ω をベクトル空間 V の開凸錐とする．開凸錐 Ω の線形自己同形群の，単位元 $e \in V$ の連結成分を G とする．もし群 G が Ω へ推移的に作用するならば，Ω は等質である（homogeneous）という．錐 Ω の双対錐 Ω^* を

$$
\Omega^* = \{y \in V \mid \langle x, y\rangle > 0, \forall x \in \bar{\Omega}\backslash\{0\}\}
$$

で定義する．ここで $\langle\, ,\rangle$ は V の内積であり，$\bar{\Omega}$ は Ω の閉包である．もし $\Omega = \Omega^*$ ならば，錐 Ω は自己双対である（self-dual）という．錐 Ω が対称である（symmetric）とは，Ω が等質かつ自己双対であることをいう．

本章最後に，ユークリッド的ジョルダン代数と対称錐との関係を述べる．ユークリッド的ジョルダン代数 V の元の平方集合 $\Omega = \{x^2 \mid x$ は可逆$\}$ は対称錐である．以下ジョルダン代数 V 上のトレース（trace）により V 上の内積 $\langle\, ,\rangle$ を定義する（i.e. $\langle x, y\rangle = tr(x * y)$．行列の場合通常のトレースに一致する.）．この内積は結合的正定値対称双線形 2 次形式であるゆえ，ジョルダン代数 V から対称錐 Ω を生成することができる．また単純ジョルダン代数 V は，ローレンツ空間・実対称行列集合・複素エルミート行列集合・四元数エルミート行列集合・八元数 3 次正方エルミート行列集合の 5 種類である．これらから生成される対称錐は順に，ローレンツ錐（2 次錐）・半正定値実対称行列集合・半正定値複素エルミート行列集合・半正定値四元数エルミート行列集合・半正定値八元数 3 次正方エルミート行列集合である．

3．対称錐上の最適化問題

線形制約をもつ対称錐最適化問題を設定し，主双対内点法とジョルダン代数とのかかわりについて述べる．主として文献 [2] に挙げられている内容である．

ベクトル空間 V をユークリッド的ジョルダン代数とし，Ω を V から生成される対称錐とする．集合 X を V のベクトル部分空間とし，$X^\perp$ を X の直交補空間（内積 $\langle x, y\rangle = tr(x * y)$ に関するもの）とする．主問題（P）と双対問題（D）の組により定められる対称錐線形計画問題を考察する．：　与えられた $a \in X$，$b \in X^\perp$ に対し

主問題（P）　：$\langle a, x\rangle \to min$, s.t. $x \in (b + X) \cap \bar{\Omega} = P$,

双対問題（D）：$\langle b, y\rangle \to min$, s.t. $y \in (a + X^\perp) \cap \bar{\Omega} = D$.

以下，$ri(P)=(b+X)\cap\Omega\neq\emptyset$, $ri(D)=(a+X^{\perp})\cap\Omega\neq\emptyset$ を仮定する（ri は相対的内点を表す.）．次の命題が成り立つ.

LEMMA 1.([2]) 正数 β に対し 2 つの最小化問題を次で与える.

$f_\beta(x)=\beta\langle a,x\rangle-\log\det(x)\to min$, s.t. $x\in ri(P)$, (P_β)

$g_\beta(y)=\beta\langle b,y\rangle-\log\det(y)\to min$, s.t. $y\in ri(D)$. (D_β)

（ジョルダン代数としての det を用いる.）

このとき $x(\beta)$ と $y(\beta)$ がそれぞれ最小化問題（P_β），（D_β）の最適解であるための必要十分条件は

$$x(\beta)\in ri(P),\ y(\beta)\in ri(D),\ x(\beta)*y(\beta)=\frac{e}{\beta}$$

である．また $\beta\to+\infty$ のとき，$x(\beta)$ と $y(\beta)$ は線形計画問題（P），(D）の最適解へそれぞれ収束する.

正数 β を無限大値へ近づけるときの $x(\beta)$ の軌跡を主問題の中心パス（central path）とよぶ．同様に $y(\beta)$ の軌跡を双対問題の中心パスとよぶ．また主問題（P）と双対問題（D）の組を実行可能領域の直積 $ri(P)\times ri(D)$ 上の計画問題とみなし，組 $(x(\beta),y(\beta))$ の軌跡を中心パスとよぶこともある.(実際には，$x(\beta)$ と $y(\beta)$ のある種の隔たりを表す量が更に 1 つ組み合わせられる.）かつ中心パスは実行可能領域の内点を通り，実行可能領域の境界の最適解へ近づく．ゆえに主問題と双対問題の組を中心パスにより求解することは主双対内点法とよばれる（[5]）．また関数 $\psi(x)=-\log\det(x)$，$\psi'(y)=-\log\det(y)$ は障壁関数とよばれ，下に狭義凸かつ境界で発散するよう定義される.

次の補題は LEMMA 1 の証明に使われる．次章以降でも利用される.

PROPOSITION 1.([2]) 最小化問題（P_β），（D_β）それぞれの最適解 $x(\beta)$，$y(\beta)$ は次をみたす．逆に次を満たす最適解は（P_β），（D_β）それぞれに対し唯一定まる.

$$x(\beta)\in ri(P),\ y(\beta)\in ri(D),\ \beta a-x(\beta)^{-1}\in X^{\perp},\ \beta b-y(\beta)^{-1}\in X,\ y(\beta)=\frac{1}{\beta}x(\beta)^{-1}$$

4．最適化問題の双対幾何構造

前章に引き続き V をユークリッド的ジョルダン代数とし，Ω を V から生成される対称錐とする．対称錐 Ω に双対幾何構造を定義した後，最適解 $x(\beta)$ と $y(\beta)$ の性質を通じて主双対内点法の双対幾何を考察する．なお対称錐上の双対幾何構造（dualistic structure）に関しては文献 [13] にも解説されている.

ベクトル空間としての V の基底を $\{e_1,\ldots,e_n\}$ とし，$x\in V$ の成分を表す関数 $x^1,\ldots,x^n$ を $x=\sum_{i=1}^n x^i(x)e_i$ で定義する．関数の組 $\{x^1,\ldots,x^n\}$ を V のアファイン座標系とみなす.(座標系と単によぶこともある．または後述の双対座標系（dual coordinate）との対比で主座標系（primal coordinate）ともよぶ.）ジョルダン代数 V 上の標準平坦アファイン接続（canonical flat affine connection）を ∇ とする．接続とは接ベクトルの平行移動を定める概念である．今の場合座標系を用いると $\nabla_{\partial/\partial x^i}\partial/\partial x^j=0$ と表せる．対称錐 Ω 上へ狭義凸関数 $\psi(x)=-\log\det(x)$ ならびにリーマン計量 $g=\nabla d\psi$ を定義

する．すると (Ω,∇,g) は平坦統計多様体（flat statistical manifold）となる．平坦とは曲率が0となることである．平坦ならば，主座標系に沿っての「直線」は ∇ についての測地線（geodesic）となる．

次に双対統計多様体（dual statistical manifold）を構成する．対称錐 Ω 上の双対座標系 $\{x'_1,\ldots,x'_n\}$ は $x'_i=-\partial\psi/\partial x^i$ により定義される．双対平坦アファイン接続（dually flat affine connection）∇' は $\nabla'_{\partial/\partial x'_i}\partial/\partial x'_j=0$ により定められる．このとき (Ω,∇',g) は平坦統計多様体となり，元の多様体 (Ω,∇,g) の双対平坦統計多様体とよばれる．双対座標系についての「直線」は ∇' についての測地線（geodesic）となる（∇ についての測地線と限らない.）．さらに (g,∇,∇') は双対平坦構造とよばれる．多様体が平坦と限らない場合も言い表すなら，同様の組は双対構造（または双対幾何構造，情報幾何構造）とよばれる．なお統計多様体の一般的定義は別途存在するが，ここで深入りしない．

第3章に挙げた対称錐線形計画問題（P）の，双対幾何構造を紹介する．実行可能領域の内点集合 $ri(P)$ へ (Ω,∇,g) を制限した統計部分多様体を $(ri(P),\nabla,g)$ と記す．また $ri(P)$ へ (Ω,∇',g) を制限した統計部分多様体は $(ri(P),\nabla',g)$ と記され，$(ri(P),\nabla,g)$ の双対多様体となる．これら2つの部分多様体は同じ集合 $ri(P)$ 上に構成されている．しかし異なる接続を与えるため，異なる統計多様体として扱う．主問題（P）の内点流 $x(\beta)$ の挙動には，$(ri(P),\nabla,g)$ を動くという捉らえ方と，$(ri(P),\nabla',g)$ を動くという捉え方がある．双対接続 ∇' に着目すれば次が従う．

THEOREM 1.　最小化問題（P_β）の最適解 $x(\beta)$ は $\beta\to+\infty$ のとき $(ri(P),\nabla',g)$ 上の測地線に沿って主問題（P）の最適解へ収束する．

Proof. 点 $x\in\Omega$ の双対座標系は，逆元 x^{-1} の主座標系から得られる．すなわち

$$x'_i=x^i\circ\iota=-\frac{\partial\psi}{\partial x^i}$$

である ([13])．ここで写像 $x\mapsto x^{-1}$ を Ω 上へ制限したものを，ι とした．写像 ι は1対1かつ上への写像となる．したがって x での双対座標系の値 $\{x'_1(x),\ldots,x'_n(x)\}$ は，同じ値をベクトル空間上の元としてもつ点 $x^{-1}\in\Omega$ と同一視される．この同一視により双対座標系を x^{-1} で表す．

さて $\{\bar{e}_1,\ldots,\bar{e}_m\}$ をベクトル部分空間 X の基底とし，$\{\bar{e}_{m+1},\ldots,\bar{e}_n\}$ をベクトル補空間 $X^\perp$ の基底とする（m は X の次元）．そして $\{\bar{x}^1,\ldots,\bar{x}^n\}$ を V のアファイン座標系とみなす．ただし $x=\sum_{i=1}^n\bar{x}^i(x)\bar{e}_i$ とする．線形領域 $b+X$ 上のアファイン座標系を $\{\bar{x}^1,\ldots,\bar{x}^m\}$ とする．これを実行可能領域 $ri(P)=(b+X)\cap\Omega$ 上へ制限したものを，$ri(P)$ 上の主座標系とする．実行可能領域 $ri(P)$ 上の双対座標系 $\{\bar{x}'_1,\ldots,\bar{x}'_m\}$ を $\bar{x}'_i=-\partial\psi/\partial\bar{x}_i$. により定義する．このとき，$\{\bar{e}_1,\ldots,\bar{e}_n\}$ が $\{e_1,\ldots,e_n\}$ の線形結合で得られることと，$x'_i=x^i\circ\iota$ とを思い出す．すると $x\in ri(P)$ に対し，x^{-1} を X へ $\langle\,,\rangle$ について射影した点の座標 $\{\bar{x}^1(x^{-1}),\ldots,\bar{x}^m(x^{-1})\}$ は，部分多様体としての双対座標 $\{\bar{x}'_1(x),\ldots,\bar{x}'_m(x)\}$ となることがわかる．一方 PROPOSITION 1 より $\beta a-x(\beta)^{-1}\in X^\perp$, $a\in X$ である．ゆえに点 $x\in ri(P)$ の $ri(P)$ 上での双対座標系の値は定ベクトルのスカラー倍 βa であ

る．したがって最小化問題（P_β）の最適解 $x(\beta)$ は，双対多様体 $(ri(P),\nabla',g)$ 上の直線に沿って収束する． □

最適解 $x(\beta)$ の軌跡を表すパラメータ β は測地線の方程式を表すパラメータに必ずしも一致しない．ゆえに測地線であると明言せず「測地線に沿う」という表現を用いる．

THEOREM 1は，ＬＰならびにＳＤＰに対する同様の結果（[1][3]）を対称錐計画法へ拡張したものである．また双対多様体 $(ri(P),\nabla',g)$ を (Ω,∇',g) の部分多様体とみなすときの第二基本形式（埋め込み曲率）が大きければ最適解の軌跡をニュートン法でたどる際の反復計算が大きくなることが，[3] で指摘されている．この事実は，本研究で扱う対称錐計画法でも同様である．

THEOREM 1の証明から導かれる性質をまとめる．点 x^{-1} の座標 $\{\bar{x}^1(x^{-1}),\ldots,\bar{x}^m(x^{-1})\}=\{\bar{x}'_1(x),\ldots,\bar{x}'_m(x)\}$ を $X^\perp$-成分とよんでいることに注意せよ．他の点や $X^\perp$-成分に関しても，同様である．

LEMMA 2.　最小化問題（P_β）の最適解 $x(\beta)\in ri(P)=(b+X)\cap\Omega$ に対し次が成り立つ．

(1) 対称錐 Ω 上における双対座標 $x(\beta)^{-1}$ の，X-成分は βa であり，$X^\perp$-成分は $\beta a - x(\beta)^{-1}$ である．

(2) 双対最小化問題（D_β）の最適解 $y(\beta)$ の $X^\perp$-成分は，Ω 上における $x(\beta)$ の双対座標の $X^\perp$-成分を $\frac{1}{\beta}$ 倍したものである．

(3) $\beta\to+\infty$ における最適解 $y(\beta)$ の軌跡は，最適解 $x(\beta)$ の (Ω,∇',g) での軌跡を線形制約領域 $ri(D)=(a+X^\perp)\cap\Omega$ 上へ，ジョルダン代数 V の原点を中心に縮小（または拡大）射影したものである．

(4) $x(\beta)^{-1}$ の座標 $\{\bar{x}^1(x^{-1}),\ldots,\bar{x}^m(x^{-1})\}=\{\bar{x}'_1(x),\ldots,\bar{x}'_m(x)\}$ の各成分はそれぞれ単調増加または単調減少のいずれかである．

5．スケーリングと作用素平均

本章ではLEMMA 1にて仮定した条件

$$x(\beta)\in ri(P),\ y(\beta)\in ri(D),\ x(\beta)*y(\beta)=\frac{e}{\beta}$$

を，主問題（P）と双対問題（D）の組のなす最適化問題に同値な最適化問題として扱う．

パラメータ $\beta\to+\infty$ に対する $(x(\beta),y(\beta))$ の極限をニュートン法で求めることが，研究されている．ニュートン法の探索方向は次の方程式をみたす $\Delta x\in X,\Delta y\in X^\perp$ として与えられる．

$$L(y)\Delta x+L(x)\Delta y=\frac{e}{\beta}-x*y,\ x\in ri(P),\ y\in ri(D)$$

実際には Ω の自己同形群 $G(\Omega)$ の元によりスケーリング（scaling）された探索方向が，しばしば設定される．スケーリングとは群作用により不変な構造を作ることである．スケーリングによる探索方向は，群作用によりある点へ引き戻せば同じ方程式をみたすようにと定義される．参考までに，対称錐上の双対接続の計算法は自己同形 $P(x)^{-\frac{1}{2}}$ による単位元への引き戻しにより各点に対し同じ方法（ジョルダン積の計算）となる．レ

ビ・チビタ接続 $(\nabla+\nabla')/2$ の計算方法についても単位元への引き戻しにより同じ方法となる．これらの事実にもスケーリングが現れている（[13]）．

たとえば次の探索方向が提案されている ([5][6]).

(1) HRVW/KSH/M 方向（Helmberg-Rendl-Vanderbei-Wolkowicz/小島-進藤-原/Monteiro direction）：主スケーリング（primal scaling）

$(\Delta x, \Delta y) := (P(y)^{-\frac{1}{2}}\Delta\tilde{x}, P(y)^{\frac{1}{2}}\Delta\tilde{y}) \in V \times V,$

where $(\tilde{x}, \tilde{y}) := (P(y)^{\frac{1}{2}}x, P(y)^{-\frac{1}{2}}y) = (P(y)^{\frac{1}{2}}x, e).$

(2) NT 方向（Nesterov-Todd direction）：主双対スケーリング（primal-dual scaling）

$(\Delta x, \Delta y) := (P(z)^{\frac{1}{2}}\Delta\tilde{x}, P(z)^{-\frac{1}{2}}\Delta\tilde{y}) \in V \times V,$

where $(\tilde{x}, \tilde{y}) := (P(z)^{-\frac{1}{2}}x, P(z)^{\frac{1}{2}}y,$ $^{\exists 1}z \in \Omega$ such that $P(z)^{-\frac{1}{2}}x = P(z)^{\frac{1}{2}}y.$

点 $x \in \Omega$ に対し，$P(x)^{-\frac{1}{2}} \in G(\Omega)$ かつ $P(x)^{-\frac{1}{2}}x = e$ であることに注意せよ．なお $\Delta\tilde{x} \in X, \Delta\tilde{y} \in X^{\perp}$ である．

NT 方向に関して次が従う．

LEMMA 3. NT 方向（主双対スケーリング）を定める自己同形 $P(z)^{-\frac{1}{2}}$ について

$$z = y^{-\frac{1}{2}}(y^{\frac{1}{2}}xy^{\frac{1}{2}})^{\frac{1}{2}}y^{-\frac{1}{2}}$$

は y と x の幾何平均である．すなわち z はレビ・チビタ接続による y から x への測地線の中点である．

行列ならびに線形作用素の幾何平均については [14] に調べられている．幾何平均の定義は対称錐上へ拡張されている．元 z が幾何平均となることは [4] で指摘されている．測地線の中点と幾何平均との対応については [15] にて指摘されている．

6．まとめ

対称錐線形計画法の主双対内点法に現れる双対幾何（情報幾何）概念について考察した．主問題の中心パスの双対座標を対称錐での双対座標として捉えれば，線形制約領域に直交する成分が双対問題の中心パスに現れることを特に指摘した．

今後の課題として次が挙げられる．

(1) 曲線 $x(\beta)^{-1}$ の (Ω,∇',g) への埋め込まれかた（埋め込み曲率など）を，LEMMA 2.(3) を利用し曲線 $y(\beta)$ の挙動から調べること．そして反復計算の回数評価へ利用すること．

(2) LEMMA 3 の利用方法を見出すこと．

参考文献

○ 最適化について

[1] K. Tanabe : Center flattening transformation and a centered Newton method for linear programing, Manuscript presented at MP seminer, The Operations Research Society of Japan, July (1987) .

[2] L. Faybusovich: Linear systems in Jordan Algebras and primal-dual Interior-point Algorithms, J. Computational and Applied Math. 86, 149-175 (1997).

[3] 小原敦美：半正定値計画問題に関する情報幾何を用いた考察，統計数理，Vol.46, No.2, 317-334 (1998).

[4] J.F.Sturm: Similarity and Other Spectral Relations for Symmetric Cones, 312,135-154 (2000).

[5] 小島政和，土谷隆，水野真治，矢部博：内点法，朝倉書店 (2001)

[6] 脇隼人：対称錐上の線形計画問題と Euclidean Jordan 代数，東京工業大学学士論文 (2002)

○ ジョルダン代数・対称錐について

[7] J.Faraut and A.Korányi: Analysis on Symmetric Cones, Oxford (1994).

○ 情報幾何とその周辺について

[8] 甘利俊一, 長岡浩司: 情報幾何の方法, 岩波講座 応用数学 12 (1993).

[9] S.Amari and H.Nagaoka, Method of information geometry, Amer. Math. Soc., Providence, Oxford University Press, Oxford (2000).

[10] A. Ohara, N. Suda and S. Amari, Dualistic Differential Geometry of Positive Definite Matrices and Its Applications to Related Problems, Linear Algebra and Its Applications, 247, 31-53 (1996).

○ 情報幾何に関連する微分幾何（アファイン微分幾何・ヘッシアン多様体・統計多様体）について

[11] 野水克己，佐々木武：アファイン微分幾何学，裳華房 (1994)

[12] 志磨裕彦：ヘッセ幾何，裳華房 (2001)

[13] K.Uohashi and A.Ohara: Jordan Algebras and Dual Affine Connections on Symmetric Cones，Positivity, Vol.8. 369/378 (2004)

○ 作用素について

[14] F.Kubo and T.Ando: Means of Positive Linear Operators, Math. Ann., 246,205-224 (1980).

[15] A.Ohara: Geodesics for Dual Connections and Means on Symmetric Cones, Integr. Eq. Oper. Theory 50, 537-548 (2004).